

WHAT IS CLAIMED IS:

1. A mass spectrometry method for identifying differences in the level of one or more analytes between two or more sample sets comprising the steps of:

5 (a) obtaining spectra for individual samples of said two or more sample sets, wherein said spectra comprise m/z-intensity pairs, wherein an m/z intensity pair comprises an m/z identifier and a signal associated with said m/z identifier,

10 (b) for each said m/z identifier of one or more m/z identifiers from said m/z intensity pairs, determining a relationship between the corresponding signals in said spectra, and

15 (c) assigning each said relationship a rank or value based on both within-sample-set and between-sample-set signal distributions, wherein said rank or value is a measure of a likelihood that said signal arises from an analyte having a different level between said two or more sample sets.

2. The method of claim 1, wherein said relationship is determined for at least 100 different m/z identifiers.

20 3. The method of claim 1, wherein said second sample set is a standard.

4. The method of claim 1 wherein each of said different m/z identifiers is deterministically specified prior to said step (b).

25 5. The method of claim 2, wherein said m/z identifiers comprise substantially all of the m/z identifiers from said spectra.

6. The method of claim 1, wherein said step (c) relies on a parametric representation of the distribution.

30 7. The method of claim 1, wherein said step (c) relies on a non-parametric representation of the distribution.

8. The method of claim 6, wherein said step (c) comprises determining the statistical significance of the difference between measures of central tendency of said distributions in light of the variability of said distributions.

5 9. The method of claim 8, wherein said central tendency is mean.

10. The method of claim 9, wherein the statistical significance is calculated using a t-test.

10 11. The method of claim 8, wherein said m/z–intensity pairs further comprises one or more index values associated with said signal and said identifier and said relationship is determined taking into account said one or more index values.

15 12. The method of claim 11, wherein the m/z-intensity pairs are aligned along the index variable(s).

13. The method of claim 12, wherein said method further comprises normalization of data prior to said step (b).

20 14. The method of claim 13, wherein signals in a set of spectra are aligned by aligning one or more landmarks, where each of said landmarks is a peak at a particular m/z identifier and at a particular set of values of index variables.

25 15. The method of claim 14, where said landmarks are found in the data by a method comprising identifying peaks that occur in all spectra in a spectra set at the same m/z identifier and at nearly the same set of index variables, optionally smoothing the intensities as a function of index variables, and using as the landmarks the set of index variable values at which the largest smoothed intensity values occur.

30 16. The method of claim 15, wherein said spectra are aligned by shifting the set of index variable values associated with each of said landmarks to the set of index variable values associated with said landmarks in some reference spectrum, and intermediate index values are assigned by interpolation.

17. The method of claim 1, wherein significant differences at a set of m/z values are grouped together as features if at least j out of k consecutive m/z identifiers have significant differences for a particular common set of index variables, where j and k are user-specified integers with j less than or equal to k .

5

18. The method of claim 17, wherein said sufficiently wide is defined by said m/z 's span being a range greater than or equal to a specified fraction of the largest m/z in the set to be grouped.

10

19. The method of claim 13, wherein said significance requires significance over at least m out of n consecutive index variable values where m and n are user-specified integers with m less than or equal to n .

15

20. The method of claim 14, wherein signals in different sets of spectra are aligned by aligning expected signals from agents specially spiked into the samples.

20

21. The method of claim 1, wherein said relationship in analyte abundance is further quantified by first calculating an integrated signal for each condition in a region containing the significant change, and then comparing the integrated signals and using the resulting relationship as indicative of relative analyte abundances.

22. The method of claim 8, wherein identified differences are grouped to indicate those putatively arising from different charge states and/or isotopes of a single analyte.

25

23. The method of claim 8, further comprising performing one or more iterations to reduce false positives.

30

24. The method of claim 23, comprising filtering said list for false positives by finding for each identified difference the index-variable shift that minimizes some measure of distance between the intensity profiles for the two conditions and determining whether the difference is still significant after said index-variable shift, then eliminating differences that are not significant after said index-variable shift.

35

25. The method of claim 13, wherein said normalization comprises, for each spectrum and each combination of index variables, finding a measure of central tendency of a

specified subset of the signals, and dividing all the intensity values by that measure of central tendency.

26. The method of claim 8, wherein at least 3 different spectra are obtained for each sample set.

27. The method of claim 26, wherein at least 5 different spectra are obtained from each sample set.

28. The method of claim 27, wherein each of said 5 different spectra is from different samples.

29. The method of claim 26, wherein said two or more sample sets are biological samples.

30. The method of claim 29, wherein said one of more analytes are peptides or metabolic by-products.

31. The method of claim 29, wherein said measurements are obtained by coupling a surface phase separation with mass spectrometry.

32. The method of claim 29, wherein said sample sets are characterized by one of more of the following: different doses of an administered agent, the presence of a disease or disorder, different types of treatment, different genetic or epigenetic attributes, or different levels of a particular disease or disorder.

33. The method of claim 29, wherein said measurements are obtained by coupling one- or multi-dimensional liquid chromatography with mass spectrometry.

34. A computer program comprising instructions on a computer readable medium for performing steps (b) and (c) of claim 1.